

Background Information: Electronic Perception Technology
July 2006

Introduction

Computer or "machine" vision is one of the oldest and most widely-researched fields in the post-industrial era, with roots going back to the dawn of television. But today, as much an art as a science, practical applications have been constrained to either extremely narrowly defined problems - such as automatically focusing cameras, or the machine assembly of pre-specified components - or to government-sponsored applications, such as mapping or spy satellites, where vast sums of money can be thrown at extremely complex and subtle technology.

The key challenge for most applications is to separate out specific objects in a scene from the background, or from one another, given only the 2-dimensional or "flat" images provided by today's camera technologies.

Both digital still cameras and video cameras utilize a common image sensing methodology in which light reflecting from objects in a three dimensional scene is projected by a lens onto a tiny, flat, semiconductor chip that contains hundreds of thousands to millions of sensing pixels arranged in a rectangular array. Each of the pixels converts light of certain wavelengths - or colors - to electronic charges, which are then "read" by adjacent electronic circuitry, and each of the pixels represents a specific, unique feature in the scene. But the only information that such devices provide is the brightness and color of the light illuminating any individual pixel at the time it was read, and the relative location of each pixel.

Perceiving specific objects or other identifiable features in a scene, given only a 2-dimensional array of a million or so tiny colored dots, is a profoundly difficult problem. Thousands of mathematical algorithms have been researched in the last 60 years that attempt such identification, but many depend upon either a complex physical setup, such as multiple cameras and controlled light sources, or upon a foreknowledge of all of the elements or motion expected in a scene. And virtually all require substantial amounts of computing power.

Canesta has made a significant breakthrough in this regard by its invention of a low cost *electronic perception technology*. The technology includes new types of chip-based image sensors, similar in size, complexity and cost to commodity-priced video camera chips, that are uniquely able to resolve the three-dimensional features of a scene relevant to specific applications. That is, in addition to the brightness of a specific color of light reflecting from nearby objects being captured by each pixel, the *distance* from the illuminated object to each pixel is also determined. Since each pixel in the sensor is illuminated by a different feature in the scene being viewed, the result is a true, three-dimensional representation. In addition, the sensors can operate at over 30 frames per second, making possible real-time applications, such as automotive sensors that can accurately judge the size, shape, and position of a passenger, security systems that identify individuals from a three-dimensional view of their face, a projection keyboard that senses finger motions on a table top, and many others. Unlike distance data derived *mathematically* from 2-D cameras though the use of geometric algorithms, distance and size are now first order parameters, *measured directly* by the sensor. This leads to a

much more straightforward interpretation of the scene, free from assumptions that are implicit in distance data derived from 2-D cameras.

The resulting geometric information acquired continuously from the scene is so rich that, depending upon the application, very often much of the computation can be performed in real time by comparatively lightweight processors embedded right into the sensor chip itself. This is in contrast to contemporary systems that can take from several seconds to several minutes to compute a tentative 3D map of a single frame.

The availability of technology of this power in such a tiny, low-cost format means that a wide variety of machines and electronic products will be able to perceive and react to nearby objects or individuals in real time through the medium of sight. The applications are endless.

What is Electronic Perception Technology?

Electronic perception technology is technology that enables machines, consumer and electronic devices, or virtually any other class of modern product to perceive and react to objects and individuals in the nearby environment in real time, particularly through the medium of "sight," utilizing low-cost, high-performance, embedded sensors and software.

What sets electronic perception technology apart from classical "computer vision" applications, for example, is that for the first time, actionable information can be developed in real time by observation of the nearby environment utilizing an ultra-low-cost sensor technology that is a size comparable to that found in nature, and as portable...

Electronic perception technology uses a four-layer model that is mirrored in the actual implementation of the technology:

SEE -> PERCEIVE -> IDENTIFY -> REACT

These are actions that happen automatically in animals and humans, but are virtually non-existent in machines, except in specialized and usually expensive applications, supported typically with extraordinary amounts of computing power.

See means the formation of an image in some medium where it is available for further processing. In animals, sight occurs in the retina of the eye, as light excites the receptors located there, and the optic nerve forwards the data on to the brain for further processing. As we saw above, in electronics - particularly video cameras - sight occurs in light-sensitive pixels arrayed on the surface of a semiconductor device. Sensed image data is forwarded via wires to some type of electronic processor or chip.

Perceive means to analyze or process an image such that specific features or objects can be ascertained. In human terms, an aboriginal culture might see letters painted on a wall, but not perceive them as letters, having no acquaintance with written language. By contrast, another culture might *perceive* them as letters, but be unable to *identify* them (for example, a westerner confronted with Chinese characters).

Identify means to recognize an object or feature such that an action or reaction is possible.

React means to actually take that action.

The goal of electronic perception technology is to make it possible for devices or applications of any complexity, from “lightweight” appliances, PDAs, cell phones, or games, to heavyweight vehicle control, airport security, or national security-class applications, to be able to perceive objects and features in the nearby environment *such that identification and action are practical and possible*.

Canesta has taken a leadership role in defining and implementing practical electronic perception technology with the development of low-cost, semiconductor-based image sensor chip technology and powerful embedded image processing software. Referencing the four-layer model, Canesta's technology - depending upon the specific application - provides *actionable perceptions or identifications* that permit third-party applications embedding Canesta's technology to *react*.

How Electronic Perception Technology Works

Electronic perception technology actually has two principal components. First are electronic perception sensor chips, and second, image processing software embedded in such chips.

Most people understand that light takes a finite time to travel between two points - that photons of light from two different stars, for example, may have started their journeys years, or even millennia apart. Since light travels essentially at a constant speed (299,792.458 kilometers per second, or about 11.8 inches in a nanosecond - a billionth of a second), if you know the time, you can calculate the distance light has traveled.

The light illuminating each individual pixel in an image sensor comes from a different feature in the scene being viewed. Canesta recognized that if you could determine the amount of time that light takes to reach each pixel from the corresponding feature in the “scene”, you then could then calculate with certainty the exact distance to that feature. In other words, you could develop a three-dimensional “relief” map of the surfaces in the scene. In three dimensions, objects previously indistinguishable from the background, for example, metaphorically “pop” out. For a broad class of applications, this proves extremely helpful in reducing the mathematical and physical complexity that has plagued computer vision applications from the start.

In Canesta's U.S. patent (#6,580,486) entitled “Methods for CMOS-compatible three-dimensional image sensing using *quantum* efficiency modulation,” Canesta describes several of its inventions for “timing” the travel time of light to a unique, new class of low-cost sensor chips.

Fundamentally, the chips work in a manner similar to RADAR (Radio Detection And Ranging), where the distance to remote objects is calculated by measuring the time it takes radio waves to make the round trip from a transmitting antenna to a reflective object (like a metal airplane) and back. In the case of these chips, however, an unobtrusive *infrared light* is transmitted instead.

The chips, which are not fooled by ambient light, then time the duration it takes this infrared light signal to reflect back to each pixel. The result is an array of “distances” - updated as often as 60 times a second or more - that provides a mathematically-accurate, dynamic “relief” map of the surfaces being imaged. The image and distance information is then handed off to a local processor running Canesta's proprietary *perception software* that further refines the 3-D representation before passing it

off to the application-layer software, which identifies objects in the scene so the vehicle or device may understand or react as appropriate.

Since electronic perception software starts with a three-dimensional view of the world, provided “for free” by the hardware, it has a substantial advantage over classical image processing software that struggles to construct three-dimensional representations using complex mathematics, and using images from multiple cameras or points of view. This significant reduction in complexity makes it possible to embed the application-independent portion of the processing software directly into the chips themselves so they may be used in the most modestly-priced applications requiring only a fraction of the space of RADAR or other systems. In addition, Canesta’s “perception layer” accounts for the remarkable ability of the technology to compute 3-dimensional image maps at over 60 frames per second - remarkable compared to existing technology that can take from several seconds to several minutes to generate a 3-dimensional representation of a single, static frame.

On top of the “perception” layer is a framework or set of Application Program Interfaces (APIs) which enable software development partners with domain expertise - such as automotive computer vision, or facial recognition - to rapidly develop electronic perception applications that identify and track objects with robustness, accuracy, and reliability unmatched by traditional 2-D computer vision systems.

Finally, with an expectation of its use in mission critical applications such as automotive, electronic perception software features tolerant, self calibrating algorithms, and is built using a layered software model that features compact code, for ease of embedding in modest applications.

Although the foregoing discussion has focused on two specific electronic perception chip designs, Canesta - with over 16 granted hardware and software patents, and 30 additional patents filed with more on the way - has substantial research and development initiatives underway that will result in future technology disclosures, product announcements, and strategic alliances well beyond what is discussed here.

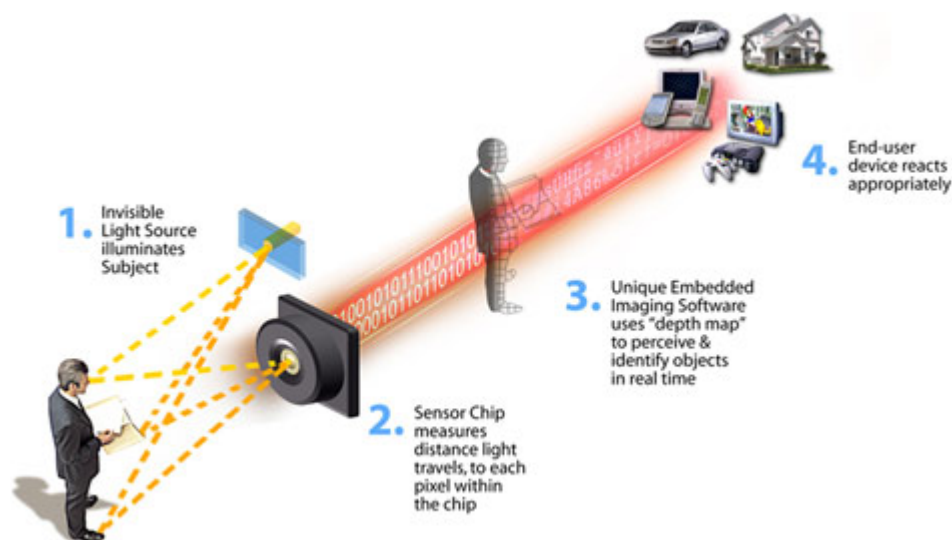


Figure 1. How Electronic Perception Technology Works

Applications for Electronic Perception Technology Case Study: Automotive

While the applications for electronic perception technology are limited only by one's imagination, the automotive industry is one of the areas with the largest need. New government regulations and consumer demands are compelling automakers to introduce new safety features at a rapid pace. If a vehicle can be made to "see", it can assist the driver in avoiding accidents, optimally deploy safety systems based on the dynamics of the crash, or even help avoid an accident altogether.

Such crucial applications, using existing or future versions of Canesta's low-cost electronic perception technology, can be identified in virtually every location within a vehicle. These include:

- *Interior* applications: occupant sensing for advanced airbag deployment, driver attention monitoring, safety "cutoff" for automatically closing windows and doors, abandoned baby and other interior security detectors, etc.;
- *Side-facing* applications: blind-spot detectors for safe lane changing, side monitors for stress-free passing in tight quarters, etc.;
- *Rear-facing* applications: rear parking assistance systems, automated rear monitors, etc.;
- *Front-facing* applications: adaptive cruise control, automated stop-and-go, follow function, pedestrian detection, etc.;
- *Whole-car* applications: systems in each of the above locations may be combined to enable the vehicle to develop full situational awareness leading to self-parking systems and "safety cocoon" 'perceptive bubble' around the vehicle.

Although other methods have been researched, prototyped, or even deployed for some of these applications, electronic perception technology has substantive advantages in many cases, particularly when:

- The vehicle itself needs to "perceive" objects or features in a scene i.e. "intelligently" process information in a scene (not just display it);
- The application is such that a foreknowledge of the scene or simplifying constraints are impractical to incorporate into the system design;
- Accuracy of recognition and tracking is critical;
- A single sensor, or at worst a single part number, is desired to simultaneously serve several different applications.

In automotive applications, compared to RADAR, electronic perception technology is less expensive and more accurate at recognizing and locating objects. While RADAR is good for determining distance to a single point or object at long range and in bad weather, it offers poor angular resolution and image fidelity, making object recognition and precise determination of location difficult. Furthermore, RADAR systems are often large in size, and have mounting limitations (such as not behind a metal bumper). Therefore, RADAR may ultimately be relegated to a role at the high end of the market due to its expense, or where the ability to operate in bad weather is essential.

Emerging automotive vision systems which use a traditional 2-D camera can certainly provide an inexpensive way to display an image for the driver to see, but the accuracy of recognition that can be

achieved is much less than what can be achieved with a 3-D image from electronic perception technology. This difference is particularly evident in unconstrained applications, where ambiguous conditions (such as "is it larger or is it closer?") frequently exist.

For example, without electronic perception technology, such ambiguous conditions are less important if the system does not need to operate correctly 100% of the time or if the application requirements can be simplified so that the general vision problem can be constrained. An example of such a simplification is a lane departure warning system; as long as there are good lane markings or a flat road with high-contrast edges, the system could operate acceptably. Outside of these simplifications, however, the system may be forced to shut down in order to avoid false warnings.

With the availability of electronic perception technology, and "true" low cost, 3-D imaging, automotive vision systems using 2-D cameras, such as those beginning to emerge today, will continue to exist in only "display only" applications, and some "easy" or highly constrained recognition applications where reduced accuracy is acceptable.

How the Technology Will Be Delivered

For automotive applications, Canesta's electronic perception technology will be delivered to Tier 1 suppliers in the form of application-specific modules which include the sensor chip, lens, light source, application program interfaces (software "hooks"), and optionally embedded application software and microprocessor. The Tier 1 then adds application-specific software, processors and/or logic, and delivers the complete, high-value-added, solution to the OEM.

The first automotive application incorporating Canesta's electronic perception technology is [expected to be occupant sensing for advanced airbag deployment.]

About Canesta

Canesta is the inventor of a revolutionary, low-cost electronic perception technology that enables machines, vehicles, and ordinary electronic devices to perceive and react to objects or individuals in real time.

When sight-enabled with Canesta's unique CanestaVision™ electronic perception chips and software, consumer, automotive, industrial, military, and medical products can gain functionality and ease of use not possible in an era when such devices were blind.

Canesta believes future applications of electronic perception technology are virtually as broad as the imagination. They may include intelligent automobile airbag systems that can sense the size and position of an occupant to control deployment and avoid injury, a low-false-alarm security system that could detect the difference between an intruder and normal activity, such as a pet moving or child visiting the bathroom at night, or robotic tools that can successfully operate in a dynamic, rather than static environment.

Canesta was founded in April 1999, and is located in San Jose, CA. The company has filed in excess of forty patents, 17 of which have been granted so far. Investment to date exceeds \$44 million, from

Carlyle Venture Partners, JP Morgan Partners, Korea Global IT Fund (KGIF), Venrock Associates and others.

####